[10] Hamilton Company, Inc., Whittier, California.

[11] R. S. Hughes Co., Los Angeles, California; Yphantis, D. A., *Ann. N. Y. Acad. Sci.*, **88**, 586 (1960).

[12] A more detailed analysis of this subject will be presented elsewhere.

[13] Fujita, H., and V. J. MacCosham, *J. Chem. Phys.*, **30**, 291 (1959).

[14] The method of moments was first used in sedimentation analysis by Baldwin in a study of boundary spreading. Baldwin, R. L., *J. Phys. Chem.*, **58**, 1081 (1954). The method has also been used by R. L. Baldwin and E. M. Shooter and by M. Meselson and G. Nazarian in analyses of sedimentation in buoyant density gradients. *Ultracentrifugal Analysis in Theory and Experiment*, ed. J. W. Williams (Academic Press, 1963), in press.

[15] Fujita, H., *J. Phys. Chem.*, **24**, 1084 (1956).

[16] Carslaw, H. S., and J. C. Jaeger, *Conduction of Heat in Solids* (Oxford: Clarendon Press, 1959), p. 259.

[17] Wayland, H., *Differential Equations Applied in Science and Engineering* (Princeton: D. Van Nostrand and Co., Inc., 1957), p. 199.

[18] Feynman, R., personal communication.

INFORMATION CONTENTS OF DISTRIBUTIONS

BY E. P. WIGNER AND MUTSUO M. YANASE*

PRINCETON UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY, PRINCETON, NEW JERSEY

Communicated April 12, 1963

1. *Introduction.*—The entropy, as usually defined, is a measure of our ignorance and, if multiplied by -1, can be considered as a measure of our knowledge of the state of a system.[1] It is a measure of our total knowledge into which the knowledge of the value of any observable enters in the same way (cf. section 3). It is this last circumstance which prompted the considerations leading to the present note. According to quantum mechanical theory, some observables can be measured much more easily than others: the observables which commute with the additive conserved quantities (energy, components of the linear and angular momenta, electric charge) can be measured with microscopic apparatuses; those which do not commute with these quantities need for their measurement macroscopic systems.[2] Hence, the problem of defining a measure of our knowledge with respect to the latter quantities arises. The present note will be restricted to the case in which there is only one conserved additive quantity; this will be denoted by k. The name "skew information" has been proposed[3] for the amount of information which an ensemble described by a state vector or a statistical matrix contains with respect to the not easily measured quantities. This information relates to the transition probabilities into states which lie askew to the characteristic vectors of the additive conserved quantities.

2. *Postulates on the Information Content.*—The requirements which an expression for the information content should satisfy are the following:

(a) If two different ensembles are united, the information content of the resulting ensemble should be smaller than the average information content of the component ensembles. By uniting two ensembles, one "forgets" from which of these a particular sample stems. Hence, the information content should decrease. Even though the present requirement is the most obvious one, it appears to be the most restrictive one and the most difficult to satisfy.

(b) The information content of the union of two systems should be the sum of the information contents of the components.

It may be well to illustrate on an example the distinction between the unions envisaged under the postulates (a) and (b). If we consider an ensemble of a system of atoms in their normal state, and another in which they are with a probability $^1/_2$ in their normal state, with a probability $^1/_2$ in their first excited state, the union of the two ensembles in the sense (a), with weights a and $1 - a$, leads to an ensemble of atoms which are with a probability $a + {}^1/_2(1 - a)$ in the normal state, with a probability $^1/_2(1 - a)$ in the first excited state. The information content I of this last ensemble should be less than $aI_1 + (1 - a)I_2$, where I_1 and I_2 are the information contents of the two initial ensembles. The union of two systems, envisaged under the present heading (b), arises if the atoms of the first ensemble and those of the second one are considered to form a single system. As a rule, one unites two systems in this way into a single one if they will interact by colliding with each other, or in some other way. If one has ensembles of the two systems, one takes a sample from the first system, and a sample from the second system, and unites these two systems to a single system which is then a sample of the ensemble of the union of the two systems. If the statistical matrices of the two ensembles are ρ_1 and ρ_2, the statistical matrix of the ensemble under (a) is $a\rho_1 + (1 - a)\rho_2$ where a and $1 - a$ are the weights of the two ensembles. The union of the two systems leads to the statistical matrix $\rho_1 \times \rho_2$ where the cross denotes the Kronecker (direct) product.

(c) The information content of an isolated system, or of an ensemble of isolated systems, should be independent of time. The change of the systems in the course of time is given in both classical and quantum mechanics by causal equations. Hence, the information which gives their state at one time gives it also at all other times as long as they are isolated.

(d) In the process which is the opposite of that considered under (b), when a joint system is separated into two parts, the information content should, in general, drop, because any knowledge of statistical correlations between the properties of the two systems will be lost by considering them separate.

(e) Finally, one should investigate the changes on the information content as a result of measurements. However, we shall not undertake this here.

3. *The Expressions for the Information Content and Their Invariances.*—In classical theory, one describes an ensemble by a distribution function $f(p_1, q_1, \ldots, p_n, q_n)$ in phase space. This gives the probability that the momenta p and coordinates q be in unit intervals at $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$, respectively. In quantum mechanics, the ensembles are described by a density matrix[4] ρ. The expectation value of a physical quantity to which the self-adjoint operator Q is coordinated, is given by the trace of $Q\rho$, to be denoted by $\mathrm{Tr}Q\rho$. The density matrix ρ is positive semi-definite and self-adjoint. The state characterized by ρ can be regarded as a mixture of orthogonal states ψ_i; these are the characteristic vectors of ρ. The characteristic value of ψ_i is the probability p_i with which ψ_i enters ρ. Since $\Sigma p_i = 1$, the trace $\mathrm{Tr}\rho = 1$ and ρ has a pure point spectrum.

The expression for the information content in classical theory is

$$I_c = \int dpdq\, f \ln f. \tag{1c}$$

$\int dpdq$ will always denote integration over the whole phase space, i.e., over all possible values of the $2n$ variables $p_1, q_1, \ldots, p_n, q_n$. In quantum mechanics,

$$I_q = \mathrm{Tr}\rho \ln \rho = \sum p_i \ln p_i. \tag{1q}$$

The last expression results by assuming ρ to be in diagonal form. Evidently, I_q is always negative, except if one deals with a pure state, that is, except if one p is 1, all others zero. Then $I_q = 0$.

The last remark shows that the information content of all pure states (states which can be described by a single state vector) is the same. This is not true for the skew information: whereas a characteristic vector of the conserved quantity k contains no such information, a state vector which lies skew to these characteristic vectors does. The expression proposed for the skew information of a pure state is[3]

$$I_s = \sum_{k\alpha} |\psi(k, \alpha)|^2 k^2 - (\sum_{k\alpha} |\psi(k, \alpha)|^2 k)^2. \tag{1s}$$

The state vector ψ in (1s) depends on the conserved quantity k and, considering that the characteristic vectors of k may be degenerate, another quantity, denoted by α. The expression (1s) is the mean square deviation of k from its average value.

The expression (1s), valid for the skew information content of a state vector, can be extended in several ways for a statistical matrix. However, the extension proposed before[3] is not tenable because it does not guarantee that condition (a) of the preceding section is fulfilled. We propose therefore

$$I_s = -{}^1/_2 \mathrm{Tr}[\sqrt{\rho}, k]^2. \tag{2}$$

The brackets denote the commutator and $\sqrt{\rho}$ the positive semidefinite square root of the positive semidefinite ρ. Since both $\sqrt{\rho}$ and k are hermitean, their commutator is skew hermitean, and the square of the commutator hermitean and negative semidefinite. Hence, because of the minus sign in (2), I_s is positive, unless ρ and k commute, in which case $I_s = 0$.

If ρ represents a pure state $\psi(k, \alpha)$, its matrix elements are, in the representation in which the coordinate axes are labeled by k and α,

$$\rho(k, \alpha;\ k', \alpha') = \psi(k, \alpha)\bar{\psi}(k', \alpha'). \tag{3}$$

The bar denotes complex conjugation. In this case, $\sqrt{\rho} = \rho$ so that the $k, \alpha;\ k', \alpha'$ matrix element of $[\sqrt{\rho}, k]$ is $\psi(k, \alpha)\bar{\psi}(k', \alpha')(k' - k)$ and one verifies that (2) contains (1s) as a special case. The purpose of the $^1/_2$ on the right of (2) is to render the two expressions numerically equal.

It is well known[5] that the expression (1c) is invariant under canonical transformations. Similarly, one easily sees that (1q) remains unchanged if one substitutes $U\rho U^{-1}$ for ρ with a unitary U. This is not true for (2), as the different directions of Hilbert space are not equivalent from the point of view of skew information. However, if U commutes with the conserved quantity k, that is, if it leaves the direction of its characteristic vectors unchanged, I for ρ and for $U\rho U^{-1}$ remain the same. In this case $k = UkU^{-1}$ and, since $(U\rho U^{-1})^{1/2} = U\sqrt{\ }U^{-1}$,

$$\begin{aligned} -{}^1/_2\mathrm{Tr}[(U\sqrt{}^{\rho}U^{-1})^{1/2}, k]^2 &= -{}^1/_2\mathrm{Tr}[U\sqrt{\rho}\,U^{-1}, UkU^{-1}]^2 \\ &= -{}^1/_2\mathrm{Tr}\{U[\sqrt{\rho}, k]U^{-1}\}^2 = -{}^1/_2\,\mathrm{Tr}U[\sqrt{\rho}, k]^2U^{-1} = -{}^1/_2\,\mathrm{Tr}[\sqrt{\rho}, k]^2. \end{aligned} \tag{4}$$

There is, evidently, one more set of transformations under which I_s is invariant: those which transform k into a constant plus k. Hence, if k is a component of the linear momentum, I_s is invariant under Galilei transformations.

4. *Verification of Condition (a) for the Proposed Expressions for the Information Content.*—As was mentioned before, in spite of its rather obvious character, this is the most difficult condition to fulfill for the expressions under consideration. Mathematically, it amounts to the inequality, to be valid for $0 \leq a \leq 1$,

$$I(a\rho_1 + (1 - a)\rho_2) \leq aI(\rho_1) + (1 - a)I(\rho_2). \tag{5}$$

In the classical case, f_1 and f_2 replace ρ_1 and ρ_2. The inequality (5) expresses the convex nature of I, as function of ρ and f, respectively. Since for $a = 0$ and $a = 1$, (5) becomes an equality, it will be guaranteed if the second derivative with respect to a of $I(a\rho_1 + (1 - a)\rho_2)$ is nonnegative. This is, then, equivalent to the condition that, for $\epsilon = 0$

$$\frac{d^2}{d\epsilon^2} I(\rho + \epsilon\sigma) \leq 0. \tag{6}$$

In (6), ρ is positive definite or at least semidefinite, σ need be only self-adjoint and such that $\rho \pm \epsilon\sigma$ be positive semidefinite for sufficiently small ϵ.

Conversely, (6) is also a necessary condition for (5) to be valid for all permissible ρ_1 and ρ_2. In order to see this, set $a = {}^1/_2$, $\rho_1 = \rho - \epsilon\sigma$, $\rho_2 = \rho + \epsilon\sigma$, with an ϵ small enough so that both ρ_1 and ρ_2 be positive semidefinite. Then, (5) becomes

$$I(\rho) \leq {}^1/_2 I(\rho - \epsilon\sigma) + {}^1/_2 I(\rho + \epsilon\sigma) \tag{7}$$

from which (6) follows if the derivatives exist and are continuous.

For the expression of the information content of classical theory, (1*c*), the inequalities (5) and (6) are well known: $f \ln f$ is a convex function of f for positive f because its second derivative, $1/f$, is positive. The convex nature of the operator function $\rho \ln \rho$ follows from the investigations of F. Krauss and of J. Bendat and S. Sherman.[6] From this, the validity of (5) for the I_q of (1*q*) follows easily. The same result was established more directly by M. Delbrück and G. Molière.[7]

We now proceed to the last case, that of skew information. In order to verify (6) for this case, we evidently need an expansion for

$$(\rho + \epsilon\sigma)^{1/2} = S + \epsilon N - \epsilon^2 T + \dots. \tag{8}$$

Since the left side is self-adjoint for all ϵ, all matrices $S, N, T, \dots$ will be self-adjoint.

Squaring both sides of (8), one obtains

$$S^2 = \rho, \quad (9) \qquad SN + NS = \sigma, \quad (10) \qquad ST + TS = N^2. \quad (11)$$

Since the positive definite or semidefinite square root of ρ occurs in (2), we shall need such a square root and S can be assumed to be positive definite or semidefinite. If it is positive definite, (10) uniquely determines N, and hence (11) determines T. One can see this most easily by assuming S to be diagonal and writing out the matrix elements of (10) and (11). If ρ and hence S are only semidefinite, the consideration becomes somewhat more involved and we shall not deal with that case. Incidentally, it follows easily from (11) and the positive definite nature of S that T is also positive semidefinite.[8] The condition (6) for I_s now reduces to the condition that the coefficient of ϵ^2 in $-\ {}^1/_2\mathrm{Tr}[S + \epsilon N - \epsilon^2 T, k]^2$ be nonnegative. Hence, we shall need

$$q = {}^1/_2\mathrm{Tr}\{[S,k][T,k] + [T,k][S,k] - [N,k]^2\} \tag{12}$$

and wish to prove that this is nonnegative.

If one writes out the matrix in the braces of (12), interchanging the factors cyclically in a suitable fashion, one obtains

$$\begin{aligned} q &= {}^1/_2\mathrm{Tr}\{2SkTk - 2kSTk + 2TkSk - 2kTSk - 2NkNk + 2kN^2k\} \\ &= \mathrm{Tr}\{SkTk + TkSk - NkNk\}. \end{aligned} \tag{13}$$

The last line follows from (11). The right side is, in terms of its matrix elements,

$$\begin{aligned} q &= \sum S_{\alpha\beta}k_{\beta\gamma}T_{\gamma\delta}k_{\delta\alpha} + T_{\alpha\beta}k_{\beta\gamma}S_{\gamma\delta}k_{\delta\alpha} - N_{\alpha\beta}k_{\beta\gamma}N_{\gamma\delta}k_{\delta\alpha} \\ &= \sum \bar{k}_{\alpha\delta}(S_{\alpha\beta}\bar{T}_{\delta\gamma} + T_{\alpha\beta}\bar{S}_{\delta\gamma} - N_{\alpha\beta}\bar{N}_{\delta\gamma})k_{\beta\gamma}. \end{aligned} \tag{14}$$

The bar denotes the conjugate complex; use has been made of the hermitean nature of all quantities.

The q in (14) is a hermitean quadratic form of the vector, the components of which are the matrix elements of k. It is the quadratic form of the matrix

$$Q = S \times \bar{T} + T \times \bar{S} - N \times \bar{N} \tag{15}$$

where the cross again denotes the Kronecker (direct) product. Hence, it must be shown that the matrix Q of (15) is positive semidefinite, if the relation (11) holds between S, T, and N, and if S is itself positive semidefinite. This is a purely mathematical theorem which has been established recently.[8] However, the proof will not be given here. Actually, it would suffice to show that Q is positive semidefinite for all vectors for the components of which the relation $k_{\alpha\beta} = \bar{k}_{\beta\alpha}$ obtains. However, the more general theorem is valid.

5. *Verification of Condition* (*b*).—It is condition (b) which prompted the use of the logarithmic function for I_c and I_q. In fact, in classical theory, the distribution function of the composite system is

$$F(p, q, p', q') = f(p, q)f'(p', q'). \tag{16c}$$

The fact that the I_c calculated with the distribution function F is equal to the sum of the two I_c calculated with the distribution functions f and f' is a matter of simple calculation.

In the quantum case, we have for the statistical matrix of the composite system

$$P = \rho \times \rho' \tag{16q,s}$$

where ρ and ρ' are the statistical matrices of the systems to be united. Hence,

$$\ln P = \ln \rho \times 1 + 1 \times \ln \rho' \tag{17q}$$

$$\sqrt{P} = \sqrt{\rho} \times \sqrt{\rho'}. \tag{17s}$$

One concludes from (17q) that $P \ln P = \rho \ln \rho \times \rho' + \rho \times \rho' \ln \rho'$, and the trace of this is

$$\mathrm{Tr}P \ln P = \mathrm{Tr}\rho \ln \rho \cdot \mathrm{Tr}\rho' + \mathrm{Tr}\rho \cdot \mathrm{Tr}\rho' \ln \rho'$$
$$= \mathrm{Tr}\rho \ln \rho + \mathrm{Tr}\rho' \ln \rho' \tag{18g}$$

since the traces of ρ and ρ' are 1. Thus, condition (b) is established for I_q. We now proceed to the consideration of I_s and denote the operator of the additive conserved quantity in the Hilbert space of the composite system by K. Since this is *additive*, $K = k \times 1 + 1 \times k'$ so that

$$[\sqrt{P}, K] = [\sqrt{\rho} \times \sqrt{\rho'}, k \times 1] + [\sqrt{\rho} \times \sqrt{\rho'}, 1 \times k']$$
$$= [\sqrt{\rho}, k] \times \sqrt{\rho'} + \sqrt{\rho} \times [\sqrt{\rho'}, k']$$

and

$$[\sqrt{P}, K]^2 = [\sqrt{\rho}, k]^2 \times \rho' + \rho \times [\sqrt{\rho'}, k']^2$$
$$+ [\sqrt{\rho}, k]\sqrt{\rho} \times \sqrt{\rho'}[\sqrt{\rho'}, k'] + \sqrt{\rho}[\sqrt{\rho}, k] \times [\sqrt{\rho'}, k']\sqrt{\rho'}.$$

However, the trace of the expressions in the second line vanishes because, for instance,

$$\mathrm{Tr}[\sqrt{\rho}, k]\sqrt{\rho} = \mathrm{Tr}\sqrt{\rho}\, k\, \sqrt{\rho} - \mathrm{Tr}k\rho = 0.$$

Hence, it follows from $\mathrm{Tr}\rho = \mathrm{Tr}\rho' = 1$

$$- {}^1/_2\mathrm{Tr}[\sqrt{P}, K]^2 = - {}^1/_2\mathrm{Tr}[\sqrt{\rho}, k]^2 - {}^1/_2\mathrm{Tr}[\sqrt{\rho'}, k']^2. \tag{18s}$$

Thus, condition (b) is valid.

6. *Verification of the Remaining Conditions.*—I could depend on time, because f and ρ depend on time. However, these changes can be represented by canonical transformations in the case of f and by a transformation $\rho \to U\rho U^{-1}$ with $U = \exp(iHt/h)$ in the case of ρ. Hence, the independence of the I on time is a special case of the invariance of these quantities, discussed at the end of section 3. This applies, in particular, also to the skew information because k is an additive *conserved* quantity; it commutes with H.

In classical theory, condition (d) is also easily verified.[9] Using the notation adopted at the beginning of the preceding section, the distribution functions for the parts into which the composite system separates become

$$f(p, q) = \int dp'dq'\, F(p, q, p', q')$$
$$f'(p', q') = \int dpdq\, F(p, q, p', q') \tag{19c}$$

where F is the distribution function of the composite system. Hence, we set

$$F(p, q, p', q') = f(p, q)f'(p', q') + g(p, q, p', q'). \tag{20c}$$

It follows from (19c) that

$$\int dpdq\, g(p, q, p', q') = \int dp'dq'\, g(p, q, p', q') = 0. \tag{21c}$$

We want to prove then that

$$\int dpdqdp'dq'\, F \ln F \geq \int dpdq\, f \ln f \times \int dp'dq'\, f' \ln f'. \tag{22c}$$

Since $F \ln F$ is a convex function for positive F (its second derivative is everywhere positive), it is everywhere larger than the first two terms of its power series. Hence,

$$F \ln F = (ff' + g) \ln (ff' + g) \geq ff' \ln ff' + (1 + \ln ff')g. \tag{23}$$

It now follows that

$$\int dpdqdp'dq' F \ln F > \int dpdqdp'dq' ff' \ln ff' + \int dpdqdp'dq'(1 + \ln ff')g. \tag{23a}$$

The first term on the right side has been calculated at the beginning of the last section: it is equal to the right side of (22*c*). Hence, (22*c*) will follow from (23*a*) if its last term vanishes. This, however, is a consequence of (21*c*), since $\ln ff' = \ln f + \ln f'$, so that one of the two equations (21*c*) applies to every term.

For the quantum theoretical case, the condition (d) was proved by Delbrück and Molière.[7] It follows also from the convex nature[6] of the operator function $P \ln P$ for positive P.

We now go over to the consideration of the skew information I_s. As was pointed out before,[3] the situation with respect to condition (d), that is, the information content of the components of a composite system, cannot be expected to be as simple in this case as in those of the standard concepts I_c and I_q. We shall consider only the case in which the composite system can be described by a wave function ψ; in this case we shall find that condition (d) can also be verified. The components of ψ will be denoted by $\psi_{\alpha\beta}$, the first index α referring to the first of the systems into which the composite system will be separated; the second, β, refers to the second such system. Hence, we have for the statistical matrices of the component systems

$$\rho_{\alpha\alpha'} = \sum_{\beta} \psi_{\alpha\beta}\bar{\psi}_{\alpha'\beta'} \qquad \rho'_{\beta\beta'} = \sum_{\alpha} \psi_{\alpha\beta}\bar{\psi}_{\alpha\beta'}. \tag{24}$$

It simplifies the formulae of the following calculation if one uses in the Hilbert space of both component systems coordinate systems in which the additive conserved quantity K is diagonal

$$K_{\alpha\beta;\alpha'\beta'} = (k_\alpha + k'_\beta)\delta_{\alpha\alpha'}\delta_{\beta\beta'}. \tag{25}$$

A further simplification results from considering ψ as a matrix, with row index α and column index β. One can then use for ψ the polar decomposition[10]

$$\psi_{\alpha\beta} = \sum_\gamma u_{\alpha\gamma}h_{\gamma\beta} \quad \text{or} \quad \psi = uh \tag{26}$$

in which u is unitary, h hermitean positive semidefinite. The decomposition (26) assumes that the indices α and β assume equally many values because both u and h are square matrices. This can be accomplished by adding rows or columns of zeros to the original matrix ψ. From the normalization condition of ψ, one infers

$$\mathrm{Tr}h^2 = \mathrm{Tr}hh^\dagger = 1 \tag{26a}$$

In terms of u and h, we have, instead of (24)

$$\rho = \psi\psi^\dagger = uh^2u^\dagger \qquad \rho' = \psi^{\mathsf{T}}\bar{\psi} = h^{\mathsf{T}}u^{\mathsf{T}}\bar{u}\bar{h} = (h^{\mathsf{T}})^2 \tag{27}$$

The dagger denotes hermitean adjoint, the $^{\mathsf{T}}$ transpose.

The skew information (I_s) of the composite system now becomes

$$I_s = \sum \bar{\psi}_{\alpha\beta}(k_\alpha + k'_\beta)^2\psi_{\alpha\beta} - \sum |\bar{\psi}_{\alpha\beta}(k_\alpha + k'_\beta)\psi_{\alpha\beta}|^2$$

$$= \mathrm{Tr}(\psi^\dagger k^2 \psi + \bar{\psi} k'^2 \psi^{\mathsf{T}} + 2\psi^\dagger k \psi k') - |\mathrm{Tr}(\psi^\dagger k \psi + \psi k' \psi^\dagger)|^2$$

$$= \mathrm{Tr}(\rho k^2 + \rho' k'^2 + \psi^\dagger k \psi k' + \psi k' \psi^\dagger k) - |\mathrm{Tr}(\psi^\dagger k \psi + \psi k' \psi^\dagger)|^2. \quad (28)$$

Some of the factors were cyclically interchanged under the trace sign.

In order to calculate the skew informations i_s and i_s' for the component systems, the positive semidefinite square roots of their statistical matrices, ρ and ρ', are needed. As (27) shows, these are $uhu^\dagger$ and h^{T} respectively. Hence,

$$i_s = \mathrm{Tr}(\rho k^2 - \sqrt{\rho} k \sqrt{\rho} k) = \mathrm{Tr}(\rho k^2 - uhu^\dagger ku\, hu^\dagger k)$$

$$i_s' = \mathrm{Tr}(\rho' k'^2 - \sqrt{\rho'} k' \sqrt{\rho'} k') = \mathrm{Tr}(\rho' k'^2 - h^{\mathsf{T}} k' h^{\mathsf{T}} k'),$$

so that

$$I_s - i_s - i_s' = \mathrm{Tr}(\psi^\dagger k \psi k' + \psi k' \psi^\dagger k + hu^\dagger ku\, hu^\dagger ku + h^{\mathsf{T}} k' h^{\mathsf{T}} k') - |\mathrm{Tr}(\psi^\dagger k \psi + \psi k' \psi^\dagger)|^2$$

$$= \mathrm{Tr}[h(u^\dagger ku + k')h(u^\dagger ku + k')] - |\mathrm{Tr}[h(u^\dagger ku + k')h]|^2. \quad (29)$$

Since h is positive semidefinite, it has a positive semidefinite square root. If the first factor h in both traces is replaced by $\sqrt{h}\,\sqrt{h}$ and one of these made the last factor, the expression for the excess of the skew information of composite over component systems assumes the form

$$I_s - i_s - i_s' = \mathrm{Tr}(jj^\dagger)\,\mathrm{Tr}(hh^\dagger) - |\mathrm{Tr}(j^\dagger h)|^2 \quad (30)$$

where $j = \sqrt{h}(u^\dagger ku + k')\sqrt{h}$. The $\mathrm{Tr}(hh^\dagger)$ could be added as a factor since (26a) shows that its value is unity. That (30) is, for arbitrary j and h, positive or zero, follows easily, however, by means of Schwarz's inequality. Hence, condition (d) is also satisfied for the I_s of (2), at least if the composite system is in a pure state. This concludes our demonstration.

We are not convinced that (2) is the only definition of the skew information which satisfies the postulates of section 2. Apart from more or less trivial generalizations of (2) (such as a linear function of I_s with positive slope), we have considered the definition $I_s = -\mathrm{Tr}([(\rho, k][\ln \rho, k])$. As Professor Dyson remarked, this and (2) are special cases of the more general expression $-\mathrm{Tr}\ [(\rho^{\vartheta}, k][\rho^{1-\vartheta}, k])$ with $0 < \vartheta \leq 1/2$, the logarithmic expression being the limiting case $\vartheta \rightarrow 0$. It has the disadvantage of giving an infinite I_s for a singular ρ (for instance, if ρ represents a pure state) unless it commutes with k. For this reason, and because of its simplicity, we prefer the I_s of (2).

One of the authors (M. M. Y.) is greatly indebted to Prof. J. R. Oppenheimer and the Institute for Advanced Study for their cordial hospitality.

* On leave from Sophia University in Tokyo.

[1] For a very brief and condensed history of the recognition of this principle, see footnote 1 of W. Weaver's article in *The Mathematical Theory of Communication* (Urbana: The University of Illinois Press, 1949), p. 45. See also the last few pages of M. v. Smoluchowski's article in *Vorträge über die kinetische Theorie der Materie und Elektrizität* (Leipzig: B. G. Teubner, 1914).

[2] Wigner, E. P., *Z. Physik*, **131**, 101 (1952); Araki, H., and M. M. Yanase, *Phys. Rev.*, **120**, 622 (1960).

[3] Wigner, E. P., *Physikertagung Wien* (Mosbach/Baden: Physik Verlag, 1962), p. 1.

[4] Landau, L., *Z. Physik*, **45**, 430 (1927); v. Neumann, J., *Nachr. Gott.*, p. 245 (1927).

[5] Gibbs, J. W., *Collected Papers* (New York: Longmans, Green and Co., 1928), p. 154; Tolman, R. C., in *The Principles of Statistical Mechanics* (Oxford University Press, 1938), p. 52.

[6] Krauss, F., *Math. Zeit.*, **41**, 18 (1936); Bendat, J., and S. Sherman, *Trans. Amer. Math. Soc.*, **79**, 58 (1955).

[7] Delbrück, M., and G. Molière, *Abh. Preuss. Akad.* (1937), p. 1.

[8] Wigner, E. P., and M. M. Yanase, to appear soon.

[9] Gibbs, J. W., *Collected Papers*, p. 159.

[10] Halmos, P. R., *Finite Dimensional Vector Spaces* (Princeton University Press, 1942), p. 138.

---

*DRUGS AFFECTING RNA AND LEARNING**

BY T. J. CHAMBERLAIN, G. H. ROTHSCHILD, AND R. W. GERARD

MENTAL HEALTH RESEARCH INSTITUTE, ANN ARBOR, MICHIGAN

Communicated April 24, 1963

The fixation or consolidation of experience upon which learning is based has been postulated to result from a dynamic process causing a permanent structural change in the neurons or neural networks of the CNS.[1, 2] Intraneuronal macromolecules, in particular ribonucleic acid (RNA), have been suggested as critical sites of these structural changes.[3] Several investigators have studied this hypothesis and have offered empirical evidence in its support.[4, 5] The following experiments were undertaken in light of this attractive hypothesis to establish whether learning phenomena can be affected by drugs which have a profound effect on RNA metabolism.

Methods and Materials.—It has been shown that a postural asymmetry in the hind limbs, induced by a unilateral cerebellar or vestibular lesion, will persist after mid-thoracic spinal cord transection, providing sufficient time is allowed for this asymmetry to "fixate" in the cord before transection.[6] The fixation time is measured between the onset of asymmetry and the transection.

Attempts were made to alter the fixation time in the spinal cord, the learning and retention of an avoidance task, and the solution of a maze problem by administering drugs reported to alter RNA metabolism: the nucleic acid (and protein) antimetabolite, 8-azaguanine; and the nucleic acid (and protein) stimulator, 1,1,3-tricyano-2-amino-1-propene, a dimer of malononitrile obtained from the Upjohn Company (U-9189).

Spinal cord fixation: A total of 124 Holtzman-derived male albino rats weighing 350–400 gm were used. Twenty were injected intraperitoneally with 50, 150, or 200 mg/kg (N = 9, 8, and 3, respectively) of 8-azaguanine, volume 1.5–4.0 cc, 1.75–6.75 (average 5) hr prior to the development of the centrally induced hind limb postural asymmetry. The lesions were made by unilateral ablation of the anterior cerebellar lobe in 3 rats and electrolytically by a stereotaxic placement of an insulated stainless steel electrode unilaterally into the vestibular nucleus in the other 17, 2 ma of anodal current being applied for 15 sec. Sixteen of the latter received electrolytic decerebrations to eliminate the necessity of any further anesthesia. Twenty-five animals were injected with 8, 10, or 15 mg/kg (N = 2, 18, and 5, respectively) of U-9189, volume 0.5–1.2 cc, 0–4 (average 2.5) hr prior to the development of the asymmetry. The lesions were made by ablation in 3 rats and electrolytically in 22. Twenty of the latter received electrolytic decerebrations. An additional 16 animals were injected with 8 mg/kg of U-9189 for 4 consecutive days, the lesioning being made